Random Variables

David Carlson

2021

Random Variables

Getting oriented

- The intellectual beginnings of probability began in gambling, and this is still the easiest way to teach it.
- In probability theory, random variables are something abstract. A random variable is a yet-to-be observed value.
- What is the probability that a coin will turn up heads? What is the probability the next card will be an ace?

- Depending on the kinds of events we are talking about, we have identified several "types" of random variables.
- These variables have known functional forms, several of which we will discuss today.
- Moreover, these functions have been extensively studied and their properties are well understood.
- The focus of this lecture is to get you familliar with these "kinds" of variables.

- Don't obsess about memorizing any of this. You will never be that far from Wikipedia. Focus on understanding:
  1. How this part of the lectures relates to the previous probabilty lecture
  2. Get a handle for the basic mapping of data types and the random variable "types" they go with (e.g., coin flips $\rightarrow$ binomial).
  3. The basic properties of random variables we care about (e.g., expected values, variance)

Levels measurement

- In empirical research, data can be classified along several dimensions. We have already distinguished between discrete (countable) and continuous (uncountable) data.
- We can also look at the precision with which the underlying quantities are measured.

Nominal

- Discrete data are nominal if there is no way to put the categories represented by the data into a meaningful order.
- Typically, this kind of data represents names (hence ‘nominal’) or attributes, like Republican or Democrat.
- A classic example would be eye color (blue, green, brown, etc.)

Ordinal

- Discrete data are ordinal if there is a logical order to the categories represented by the data, but there is no common scale for differences between adjacent categories.
- Ideology is commonly measured as ordinal: Very Liberal, Libera, Somewhat liberal, Somewhat conservative, etc.

Interval

- Discrete or continuous data are interval if there is an order to the values and there is a common scale, so that differences between two values have substantive meanings.
- Dates are a common interval data
- Year is discrete interval data

Ratio

- Discrete or continuous data are ratio if the data have the characteristics of interval data and zero is a meaningful quantity.
- This allows us to consider the ratio of two values as well as difference between them.
- Quantities measured in dollars, such as per capita GDP, are ratio data.

Types of distribution functions

You will work primarily with three types of distribution functions:

1. Probability mass functions
2. Probability density functions
3. Cumulative distribution functions

Probability Mass Functions

- The PMF of a discrete distribution maps $p : X \to [0, 1]$

Probability Mass Functions

- The PMF of a discrete distribution maps $p : X \to [0, 1]$
- Domain: possible discrete values of X, often $\subseteq \mathbb{N}$

Probability Mass Functions

- The PMF of a discrete distribution maps $p : X \to [0, 1]$
- Domain: possible discrete values of X, often $\subseteq \mathbb{N}$
- Range: $0 \leq p(x) \leq 1$

Probability Mass Functions

- The PMF of a discrete distribution maps $p : X \to [0, 1]$
- Domain: possible discrete values of X, often $\subseteq \mathbb{N}$
- Range: $0 \leq p(x) \leq 1$
- From Axioms:

$$\sum p(x) = 1$$

Probability Mass Functions

- The PMF of a discrete distribution maps $p : X \to [0, 1]$
- Domain: possible discrete values of X, often $\subseteq \mathbb{N}$
- Range: $0 \leq p(x) \leq 1$
- From Axioms:

$$\sum p(x) = 1$$

- Write

$$p(X = x|\text{parameters})$$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$
- PDF also called the "marginal dist'n of x"

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$
- PDF also called the "marginal dist'n of x"
- Domain: possible values of X, $\subseteq \mathbb{R}^1$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \rightarrow [0, 1]$
- PDF also called the "marginal dist'n of x"
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq p(x) \leq 1$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$
- PDF also called the "marginal dist'n of x"
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq p(x) \leq 1$
- From Axioms:

$$\int_{-\infty}^{\infty} p(x) = 1$$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$
- PDF also called the "marginal dist'n of x"
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq p(x) \leq 1$
- From Axioms:
$$\int_{-\infty}^{\infty} p(x) = 1$$
- Write
$$p(X = x|\text{parameters})$$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$
- PDF also called the "marginal dist'n of x"
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq p(x) \leq 1$
- From Axioms:

$$\int_{-\infty}^{\infty} p(x) = 1$$

- Write

$$p(X = x | \mathrm{parameters})$$

- For joint distn's, $p(x, y) = p(X = x, Y = y)$

Probability Density Functions

- PDF of a continuous(/discrete) distribution maps $p : X \to [0, 1]$
- PDF also called the "marginal dist'n of x"
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq p(x) \leq 1$
- From Axioms:

$$\int_{-\infty}^{\infty} p(x) = 1$$

- Write

$$p(X = x | \text{parameters})$$

- For joint distn's, $p(x, y) = p(X = x, Y = y)$
- Generally, the joint dist'n is **not** the product of the marginals

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq F(x) \leq 1$

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq F(x) \leq 1$
- Discrete dist'n: $F(x) = p(X \leq x) = \sum_{x \leq z} p(z) = \sum_{x=0}^{z} p(z)$

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq F(x) \leq 1$
- Discrete dist'n: $F(x) = p(X \leq x) = \sum_{x \leq z} p(z) = \sum_{x=0}^{z} p(z)$
- Continuous dist'n: $F(x) = p(X \leq x) = \int_{-\infty}^{x} p(z)dz$

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq F(x) \leq 1$
- Discrete dist'n: $F(x) = p(X \leq x) = \sum_{x \leq z} p(z) = \sum_{x=0}^{z} p(z)$
- Continuous dist'n: $F(x) = p(X \leq x) = \int_{-\infty}^{x} p(z)dz$
- Note: In both cases, $F(max(x)) = 1$

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq F(x) \leq 1$
- Discrete dist'n: $F(x) = p(X \leq x) = \sum_{x \leq z} p(z) = \sum_{x=0}^{z} p(z)$
- Continuous dist'n: $F(x) = p(X \leq x) = \int_{-\infty}^{x} p(z)dz$
- Note: In both cases, $F(max(x)) = 1$
- $F^{-1}(x)$ is the *quantile function* of X

(Cumulative) Distribution Functions

- The CDF of a distribution maps $F : X \to [0, 1]$
- Domain: possible values of X, $\subseteq \mathbb{R}^1$
- Range: $0 \leq F(x) \leq 1$
- Discrete dist'n: $F(x) = p(X \leq x) = \sum_{x \leq z} p(z) = \sum_{x=0}^{z} p(z)$
- Continuous dist'n: $F(x) = p(X \leq x) = \int_{-\infty}^{x} p(z)dz$
- Note: In both cases, $F(max(x)) = 1$
- $F^{-1}(x)$ is the *quantile function* of X

Discret distributions

- **Random Variable**: A random variable is a real-valued function defined on the sample space S.
- It assigns a real number to every outcome $s \in S$.
- **Discrete Random Variable**: Y is a discrete random variable if it can assume only a finite or countably infinite number of distinct values.
- Examples: number of wars per year, heads or tails, voting Republican or Democrat, number on a rolled die.

Probability Mass Function

- For a discrete random variable Y, the probability mass function (pmf) $p(x) = \Pr(X = x)$ assigns probabilities to a countable number of distinct x values such that

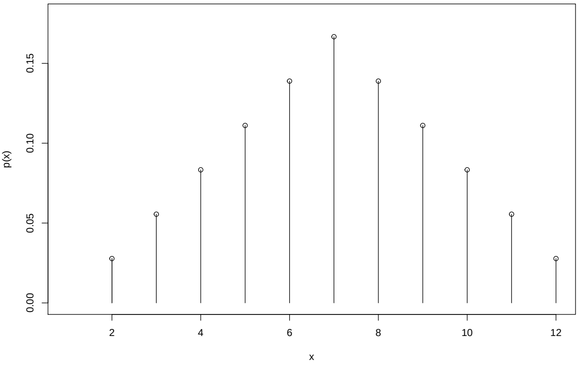
1. $0 \leq p(x) \leq 1$
2. $\sum_{x} p(x) = 1$

Example

- For one fair six-sided die, there is an equal probability of rolling any number.
- Since there are six sides, the probability mass function is then $p(y) = 1/6$ for $y = 1, \ldots, 6$.
- Each $p(y)$ is between 0 and 1.
- And, the sum of the $p(y)$'s is 1.

- If there are two six-sided dice, the probability mass funciton is shown below.

```
y<-c(1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1)/36; x<-c(2:12)
plot(x, y, xlim=c(1, 12), ylim=c(0, .18),
     xlab="x", ylab="p(x)")
segments(x0=x, y0=rep(0,12), x1=x, y1=y)
```



Cumulative distribution

- The cumulative distribution $F(x)$ or $\Pr(X \leq x)$ is the probability that Y is less than or equal to some value y, or

$$\Pr(X \leq x) = \sum_{i \leq x} p(i)$$

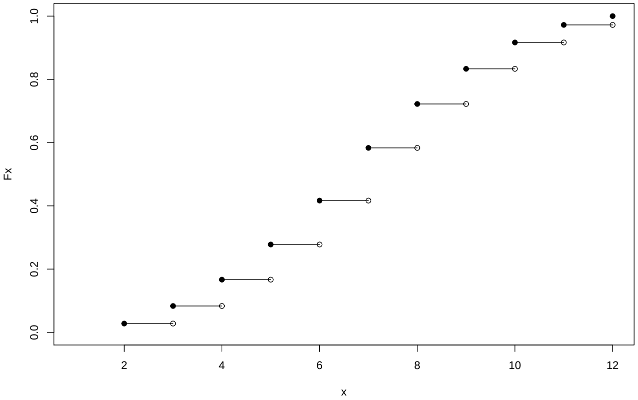
. The CDF must satisfy these properties:

  1. $F(x)$ is non-decreasing in x.
  2. $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$
  3. $F(x)$ is right-continuous.

- Example: For a fair die, $\Pr(Y \leq 1) = 1/6$, $\Pr(Y \leq 3) = 1/2$, and $\Pr(Y \leq 6) = 1$.

Example: Two Fair die

```
fx<-c(1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1)/36; x<-c(2:12)
Fx<-sapply(1:11, function(i, fx) sum(fx[1:i]), fx=fx)
plot(x, Fx, xlim=c(1, 12), ylim=c(0, 1), pch=19)
points(3:12, Fx[-11], xlim=c(1, 12), ylim=c(0, 1), pch=1)
segments(x0=2:11, x1=3:12, y0=Fx[-11], y1=Fx[-11])
```



Bernoulli Distribution

- Single trial, binary outcome $0, 1$.

Bernoulli Distribution

- Single trial, binary outcome $0, 1$.
- "Prob of success in 1 trial?"

Bernoulli Distribution

- Single trial, binary outcome $0, 1$.
- "Prob of success in 1 trial?", $x \in \{0, 1\}$
- $X \sim Bern(p)$

Bernoulli Distribution

- Single trial, binary outcome $0, 1$.
- "Prob of success in 1 trial?", $x \in \{0, 1\}$
- $X \sim Bern(p)$
- $p(X = x|p) = p^x(1-p)^{1-x}$

Bernoulli Distribution

- Single trial, binary outcome $0, 1$.
- "Prob of success in 1 trial?", $x \in \{0, 1\}$
- $X \sim Bern(p)$
- $p(X = x|p) = p^x(1-p)^{1-x}$

Political science examples

- Let $X = \begin{cases} 1 & \text{if you turnout} \\ 0 & \text{if you abstain} \end{cases}$.
  - Then, $p(X = 1|p = .4) = .4$ prob of you turning out to vote in next election, given underlying true prob $p = .4$
  - $p(X = 0|p = .4) = .6$ prob of you abstaining in next election.
- What is the probability of a of US-NKorea conflict in 2022?

Binomial distribution

- *n* $\{$*independent*$\}$, $\{$*identically distributed*$\}$ *(iid) trials, binary outcome* $0, 1$.

Binomial distribution

- *n* $\{$*independent*$\}$, $\{$*identically distributed*$\}$ *(iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*

Binomial distribution

- n {*independent*}, {*identically distributed*} *(iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*, $k \in \{0, \ldots, n\}$

Binomial distribution

- *n {independent}, {identically distributed} (iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*, $k \in \{0, \ldots, n\}$
- $X \sim Bin(n, p)$

Binomial distribution

- *n {independent}, {identically distributed} (iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*, $k \in \{0, \dots, n\}$
- $X \sim Bin(n, p)$
- $p(X = k|n, p) = \binom{n}{k} p^k (1-p)^{n-k}$

Binomial distribution

- n {*independent*}, {*identically distributed*} *(iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*, $k \in \{0, \dots, n\}$
- $X \sim Bin(n, p)$
- $p(X = k|n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$
- $X \sim Bin(1, p) \sim p(X = k|1, p) \sim Bern(p)$

Binomial distribution

- n {*independent*}, {*identically distributed*} *(iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*, $k \in \{0, \dots, n\}$
- $X \sim Bin(n, p)$
- $p(X = k|n, p) = \binom{n}{k} p^k (1-p)^{n-k}$
- $X \sim Bin(1, p) \sim p(X = k|1, p) \sim Bern(p)$
- $X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p)$, $X_1 \perp\!\!\!\perp X_2$, *then* $X_1 + X_2 \sim Bin(n_1 + n_2, p)$

Binomial distribution

- *n* {*independent*}, {*identically distributed*} *(iid) trials, binary outcome* $0, 1$.
- *"Prob of k successes in n trials?"*, $k \in \{0, \dots, n\}$
- $X \sim Bin(n, p)$
- $p(X = k|n, p) = \binom{n}{k} p^k (1 - p)^{n-k}$
- $X \sim Bin(1, p) \sim p(X = k|1, p) \sim Bern(p)$
- $X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p)$, $X_1 \perp\!\!\!\perp X_2$, *then* $X_1 + X_2 \sim Bin(n_1 + n_2, p)$
- $F(x) = \sum_{i=0}^{\lfloor x \rfloor} \binom{n}{i} p^i (1 - p)^{n-i}$

Binomial distribution

- *n* $\{$*independent*$\}$, $\{$*identically distributed*$\}$ *(iid) trials, binary outcome* $0, 1$.
- *"Prob of* k *successes in* n *trials?",* $k \in \{0, \ldots, n\}$
- $X \sim Bin(n, p)$
- $p(X = k | n, p) = \binom{n}{k} p^k (1-p)^{n-k}$
- $X \sim Bin(1, p) \sim p(X = k | 1, p) \sim Bern(p)$
- $X_1 \sim Bin(n_1, p)$, $X_2 \sim Bin(n_2, p)$, $X_1 \perp\!\!\!\perp X_2$, *then* $X_1 + X_2 \sim Bin(n_1 + n_2, p)$
- $F(x) = \sum\limits_{i=0}^{\lfloor x \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$
- *Sum of* n *Bernoullis*

Political examples:

- Prob 3 of 6 opposing Senators support an amendment:

Political examples:

- Prob 3 of 6 opposing Senators support an amendment: $p(X = 3|n = 6, p = .3) =$

Political examples:

- Prob 3 of 6 opposing Senators support an amendment: $p(X = 3|n = 6, p = .3) =$ `dbinom(3,6,prob=.3)` $\approx .19$

Political examples:

- Prob 3 of 6 opposing Senators support an amendment: $p(X = 3|n = 6, p = .3) =$ `dbinom(3,6,prob=.3)` $\approx .19$
- Prob ≥ 3 of 6 opposing Senators support an amendment:

Political examples:

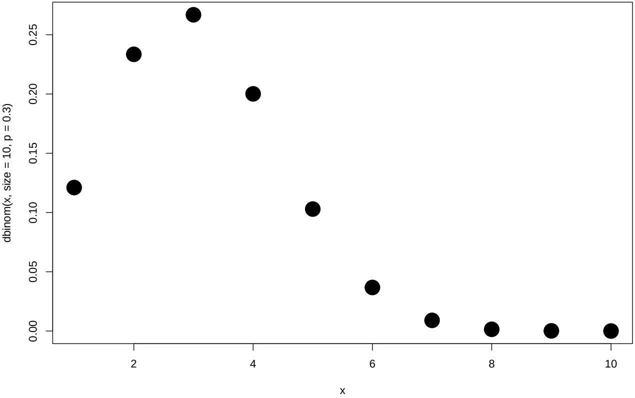
- Prob 3 of 6 opposing Senators support an amendment: $p(X = 3|n = 6, p = .3) =$ `dbinom(3,6,prob=.3)` $\approx .19$
- Prob ≥ 3 of 6 opposing Senators support an amendment: $p(X \geq 3|n = 6, p = .3) =$

Political examples:

- Prob 3 of 6 opposing Senators support an amendment: $p(X = 3|n = 6, p = .3) =$ `dbinom(3,6,prob=.3)` $\approx .19$
- Prob ≥ 3 of 6 opposing Senators support an amendment: $p(X \geq 3|n = 6, p = .3) =$ `1-pbinom(2,6,prob=.3)` $=$ `pbinom(2,6,prob=.3,lower.tail=FALSE)` $\approx .26$

Binomial PMF: N=10, p=0.3

```
x<-1:10
plot(x, dbinom(x, size=10, p=.3), cex=3, pch=19)
```



Poisson Distribution

- iid Bernoulli trial at every instant

Poisson Distribution

- iid Bernoulli trial at every instant
- “Prob of k successes in unit time?”

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$
- $\lambda > 0$, expected number of successes/events in unit time

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$
- $\lambda > 0$, expected number of successes/events in unit time
- $X \sim Pois(\lambda)$

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$
- $\lambda > 0$, expected number of successes/events in unit time
- $X \sim Pois(\lambda)$
- $p(X = k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$
- $\lambda > 0$, expected number of successes/events in unit time
- $X \sim Pois(\lambda)$
- $p(X = k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$
- $X_1 \sim Pois(\lambda_1)$, $X_2 \sim Pois(\lambda_2)$, $X_1 \perp\!\!\!\perp X_2$, then $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$
- $\lambda > 0$, expected number of successes/events in unit time
- $X \sim Pois(\lambda)$
- $p(X = k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$
- $X_1 \sim Pois(\lambda_1)$, $X_2 \sim Pois(\lambda_2)$, $X_1 \perp\!\!\!\perp X_2$, then $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$
- $F(x|\lambda) = \sum_{z=0}^{x} \frac{e^{-\lambda}\lambda^z}{z!}$

Poisson Distribution

- iid Bernoulli trial at every instant
- "Prob of k successes in unit time?", $k \in \mathbb{N}$
- $\lambda > 0$, expected number of successes/events in unit time
- $X \sim Pois(\lambda)$
- $p(X = k|\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$
- $X_1 \sim Pois(\lambda_1)$, $X_2 \sim Pois(\lambda_2)$, $X_1 \perp\!\!\!\perp X_2$, then $X_1 + X_2 \sim Pois(\lambda_1 + \lambda_2)$
- $F(x|\lambda) = \sum_{z=0}^{x} \frac{e^{-\lambda}\lambda^z}{z!}$

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4:

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X = 2|\lambda = 4) =$

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X = 2|\lambda = 4) = \texttt{dpois(2,4)}$

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X=2|\lambda=4) =$ `dpois(2,4)` $\approx .15$

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X = 2|\lambda = 4) = \texttt{dpois(2,4)} \approx .15$
  - prob of ≥ 2 HoR censures this Congress, when usually 4:

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X = 2|\lambda = 4) = \texttt{dpois(2,4)} \approx .15$
  - prob of $\geq$ 2 HoR censures this Congress, when usually 4: $p(X \geq 2|\lambda = 4) =$

Political examples

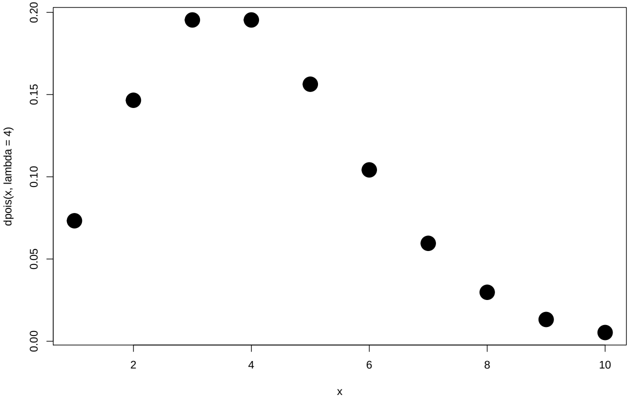
- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X = 2|\lambda = 4) = \texttt{dpois(2,4)} \approx .15$
  - prob of ≥ 2 HoR censures this Congress, when usually 4: $p(X \geq 2|\lambda = 4) = \texttt{1-ppois(1,4)}$

Political examples

- Political examples:
  - prob of 2 HoR censures this Congress, when usually 4: $p(X = 2|\lambda = 4) = \texttt{dpois(2,4)} \approx .15$
  - prob of ≥ 2 HoR censures this Congress, when usually 4: $p(X \geq 2|\lambda = 4) = \texttt{1-ppois(1,4)} \approx .91$

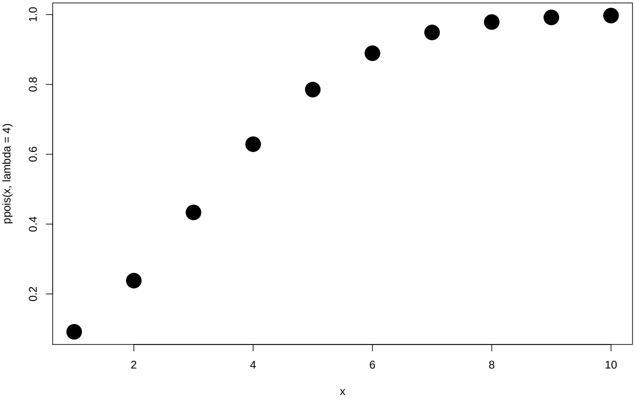
PMF for Poisson with $\lambda = 4$

```
x<-1:10
plot(x, dpois(x, lambda=4), cex=3, pch=19)
```



CDF for Poisson with $\lambda = 4$

```
x<-1:10
plot(x, ppois(x, lambda=4), cex=3, pch=19)
```



Negative Binomial Distribution

- Waiting: iid binary trials

Negative Binomial Distribution

- Waiting: iid binary trials
- "Prob of k failures before n^{th} success?"

Negative Binomial Distribution

- Waiting: iid binary trials
- "Prob of k failures before n^{th} success?", $k \in \mathbb{N}$

Negative Binomial Distribution

- Waiting: iid binary trials
- "Prob of k failures before n^{th} success?", $k \in \mathbb{N}$
- $X \sim NegBin(n, p)$

Negative Binomial Distribution

- Waiting: iid binary trials
- "Prob of k failures before n^{th} success?", $k \in \mathbb{N}$
- $X \sim NegBin(n, p)$
- $p(X = k | n, p) = \binom{n+k-1}{k} p^n (1-p)^k$

Political example:

- prob 4 days pass before 3rd roadside bomb, if each day $p = .5$:

Political example:

- prob 4 days pass before 3rd roadside bomb, if each day $p = .5$: $p(X = 4|n = 3, p = .5) =$

Political example:

- prob 4 days pass before 3rd roadside bomb, if each day $p = .5$: $p(X = 4|n = 3, p = .5) =$ `dnbinom(4,3,p=.5)` $\approx .12$

Political example:

- prob 4 days pass before 3rd roadside bomb, if each day $p = .5$: $p(X = 4|n = 3, p = .5) =$ `dnbinom(4,3,p=.5)` $\approx .12$
- prob ≤ 4 days pass before 3^{rd} roadside bomb, if each day $p = .5$:

Political example:

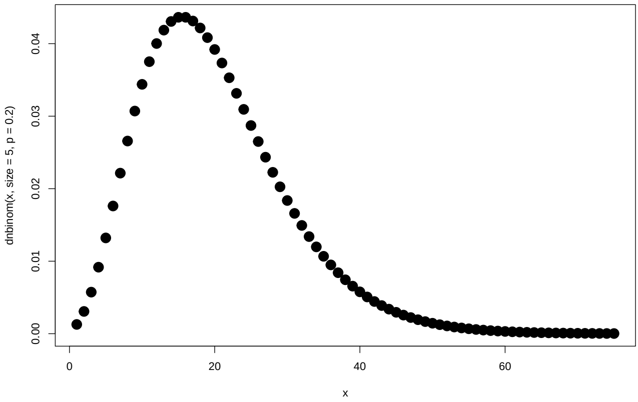
- prob 4 days pass before 3rd roadside bomb, if each day $p = .5$: $p(X = 4|n = 3, p = .5) =$ `dnbinom(4,3,p=.5)` $\approx .12$
- prob ≤ 4 days pass before 3^{rd} roadside bomb, if each day $p = .5$: $p(X \leq 4|n = 3, p = .5) =$

Political example:

- prob 4 days pass before 3rd roadside bomb, if each day $p = .5$: $p(X = 4|n = 3, p = .5) =$ `dnbinom(4,3,p=.5)` $\approx .12$
- prob ≤ 4 days pass before 3^{rd} roadside bomb, if each day $p = .5$: $p(X \leq 4|n = 3, p = .5) =$ `pnbinom(4,3,p=.5)` $\approx .77$

PMF for Negative Binomial with n=5, p=.2

```
x<-1:75
plot(x, dnbinom(x, size=5, p=.2), cex=2, pch=19)
```



Other distributions you might encounter

- Geometric
- HyperGeometric
- Multinomial (The dice examples)

Continuous Random Variables:

- X is a continuous random variable if there exists a nonnegative function $f(x)$ defined for all real $y \in (-\infty, \infty)$, such that for any interval A,

$$\Pr(x \in A) = \int_A f(x)dx$$

- Examples: income, GNP, temperature

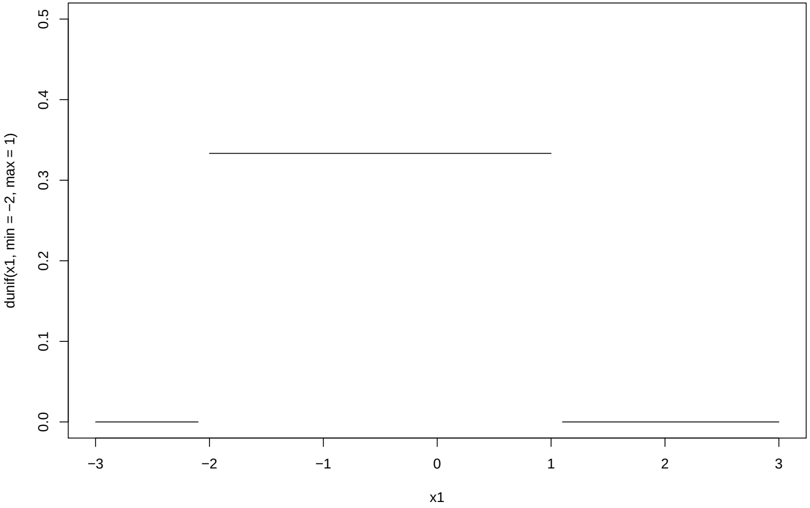
Probability Density function

- The function f above is called the probability density function (pdf) of x and must satisfy
  1. $f(x) \geq 0$
  2. $\int\limits_{-\infty}^{\infty} f(x)dx = 1$
- Note also that $\Pr(X = x) = 0$ — i.e., the probability of any point x is zero.

Example: Uniform distribution

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$$

PDF for Uniform(-2, 1)



Cumulative Distribution

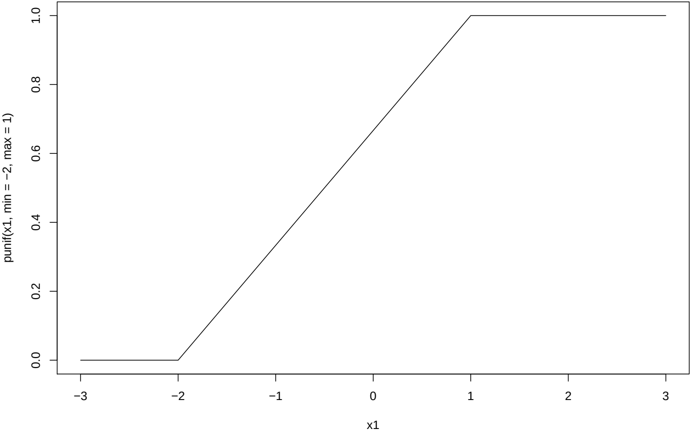
- Because the probability that a continuous random variable will assume any particular value is zero, we can only make statements about the probability of a continuous random variable being within an interval.
- The cumulative distribution gives the probability that X lies on the interval $(-\infty, x)$ and is defined as

$$F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} f(s)ds$$

- Note that $F(x)$ has similar properties with continuous distributions as it does with discrete
  - non-decreasing, continuous (not just right-continuous),
  - and $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

CDF for Uniform(-2, 1)

```
x1<-seq(-3, 3, by=.1)
plot(x1, punif(x1, min=-2, max=1), cex=2, pch=19,
     type="l", xlim=c(-3,3), ylim=c(0, 1))
```



- Similarly, we can also make probability statements about X falling in an interval $a \leq x \leq b$.

$$\Pr(a \leq x \leq b) = \int\limits_a^b f(x) dy$$

Example: $f(x) = 1, \quad 0 < x < 1$. Find $F(x)$ and $\Pr(.5 < y < .75)$.

$$F(y) = \int\limits_0^y f(s) ds = \int_0^y 1 ds = s|_0^y = y$$

$$\Pr(.5 < y < .75) = \int_{.5}^{.75} 1 ds = s|_{.5}^{.75} = .25$$

- Finally, note that:

$$F'(y) = \frac{dF(y)}{dy} = f(y)$$

Uniform Distribution

- $X \sim Unif(a, b)$

Uniform Distribution

- $X \sim Unif(a, b)$
- $x \in [a, b]$

Uniform Distribution

- $X \sim Unif(a, b)$
- $x \in [a, b]$
- $p(x) = \frac{1}{b-a}$

Uniform Distribution

- $X \sim Unif(a, b)$
- $x \in [a, b]$
- $p(x) = \frac{1}{b-a}$
- $F(x) = \begin{cases} 0 & \text{for} \quad x < a \\ \frac{x-a}{b-a} & \text{for} \quad a \leq x < b \\ 1 & \text{for} \quad x \geq b \end{cases}$

Uniform Distribution

- $X \sim Unif(a, b)$
- $x \in [a, b]$
- $p(x) = \frac{1}{b-a}$
- $F(x) = \begin{cases} 0 & \text{for} \quad x < a \\ \frac{x-a}{b-a} & \text{for} \quad a \leq x < b \\ 1 & \text{for} \quad x \geq b \end{cases}$
- Common: $X \sim Unif(0, 1)$

Uniform Distribution

- $X \sim Unif(a, b)$
- $x \in [a, b]$
- $p(x) = \frac{1}{b-a}$
- $F(x) = \begin{cases} 0 & \text{for} \quad x < a \\ \frac{x-a}{b-a} & \text{for} \quad a \leq x < b \\ 1 & \text{for} \quad x \geq b \end{cases}$
- Common: $X \sim Unif(0, 1)$

Political examples:

- "Suppose voter's probability of turnout is draw from uniform"
- Suppose the ideology of an agent is drawn from the uniform distribution

Exponential Distribution

- $X \sim Expo(\beta)$

Exponential Distribution

- $X \sim Expo(\beta)$
- $x \geq 0$

Exponential Distribution

- $X \sim Expo(\beta)$
- $x \geq 0$
- $\beta > 0$, $\beta =$ mean duration, "scale" parameter

Exponential Distribution

- $X \sim Expo(\beta)$
- $x \geq 0$
- $\beta > 0$, $\beta =$ mean duration, "scale" parameter
- $p(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$

Exponential Distribution

- $X \sim Expo(\beta)$
- $x \geq 0$
- $\beta > 0$, $\beta =$ mean duration, "scale" parameter
- $p(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}$

χ^2 Distribution

- $X \sim \chi_n^2$

χ^2 Distribution

- $X \sim \chi_n^2$
- $x > 0$

χ^2 Distribution

- $X \sim \chi^2_n$
- $x > 0$
- $n \in \mathbb{Z}_+$

χ^2 Distribution

- $X \sim \chi^2_n$
- $x > 0$
- $n \in \mathbb{Z}_+$ 0 $p(x) = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$

χ^2 Distribution

- $X \sim \chi^2_n$
- $x > 0$
- $n \in \mathbb{Z}_+$ 0 $p(x) = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$
- n sometimes called "degrees of freedom"
- $X_1 \sim \chi^2_{n_1}$, $X_2 \sim \chi^2_{n_2}$, $X_1 \perp\!\!\!\perp X_2$, then $X_1 + X_2 \sim \chi^2_{n_1+n_2}$

χ^2 Distribution

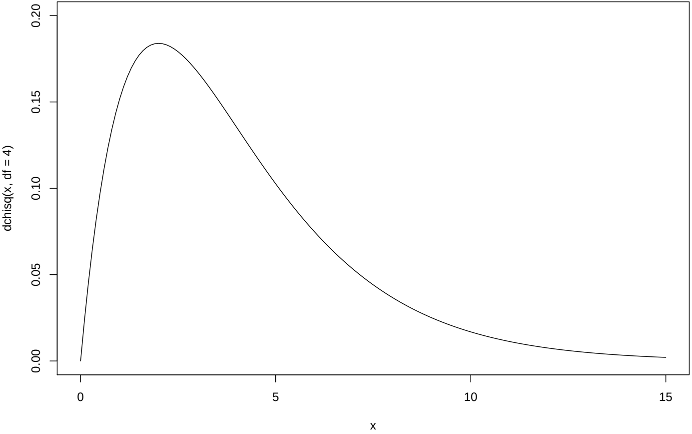
- $X \sim \chi^2_n$
- $x > 0$
- $n \in \mathbb{Z}_+$ 0 $p(x) = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$
- n sometimes called "degrees of freedom"
- $X_1 \sim \chi^2_{n_1}$, $X_2 \sim \chi^2_{n_2}$, $X_1 \perp\!\!\!\perp X_2$, then $X_1 + X_2 \sim \chi^2_{n_1+n_2}$
- $\chi^2_n \sim Gamma\left(\frac{n}{2}, 2\right)$

χ^2 Distribution

- $X \sim \chi^2_n$
- $x > 0$
- $n \in \mathbb{Z}_+$ 0 $p(x) = \frac{1}{\Gamma(\frac{n}{2})2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$
- n sometimes called “degrees of freedom”
- $X_1 \sim \chi^2_{n_1}$, $X_2 \sim \chi^2_{n_2}$, $X_1 \perp\!\!\!\perp X_2$, then $X_1 + X_2 \sim \chi^2_{n_1+n_2}$
- $\chi^2_n \sim Gamma\left(\frac{n}{2}, 2\right)$

PDF of χ^2 distribution with $n = 4$

```
x<-seq(0, 15, by=.1)
plot(x, dchisq(x, df=4), cex=2, pch=19,
     type="l", xlim=c(0,15), ylim=c(0, .2))
```



Political examples:

- Model relationships between table rows/columns
  $X^2 = \sum_{i=1}^{r} \sum_{i=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}$
- regression statistics

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$
- $\sigma > 0$

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$
- $\sigma > 0$
- $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$
- $\sigma > 0$
- $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Common: $X \sim N(0, 1)$, "standard normal"

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$
- $\sigma > 0$
- $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Common: $X \sim N(0, 1)$, "standard normal"
- $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Normal (Gaussian) Distribution

- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$
- $\sigma > 0$
- $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Common: $X \sim N(0, 1)$, "standard normal"
- $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- $\Phi(x) =$ standard normal CDF – that no one knows

Normal (Gaussian) Distribution

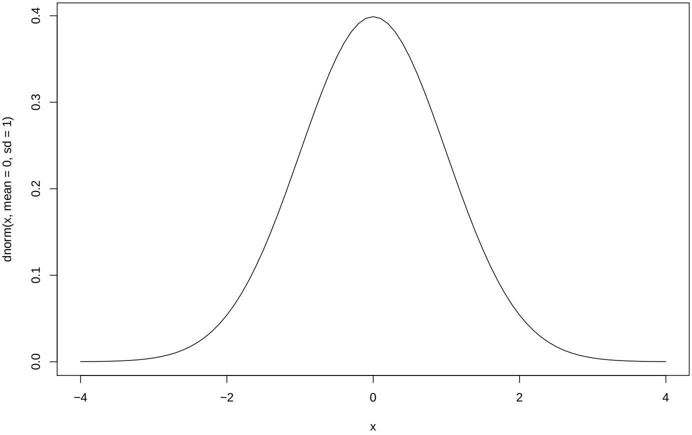
- $X \sim N(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $\mu \in \mathbb{R}$
- $\sigma > 0$
- $p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
- Common: $X \sim N(0, 1)$, "standard normal"
- $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$
- $\Phi(x) =$ standard normal CDF – that no one knows

Political examples:

- population quantities, asymptotic/known variance sampling distributions
- $\Phi(x) = p(X = 1)$ is the basic *probit model*

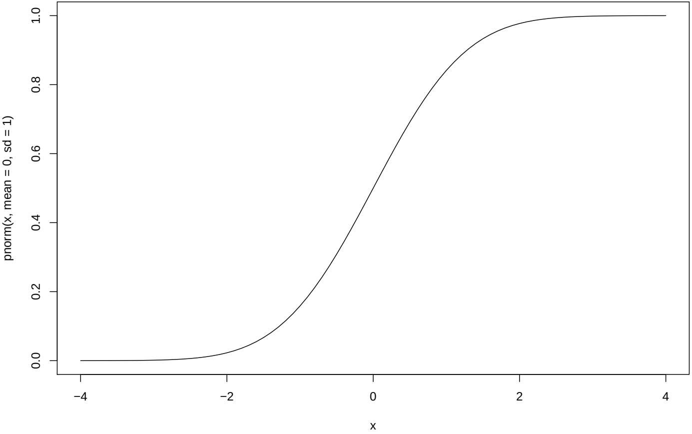
PDF for standard normal distribution

```
x<-seq(-4, 4, by=.1)
plot(x, dnorm(x, mean=0, sd=1),  type="l")
```



CDF for standard normal distribution

```
x<-seq(-4, 4, by=.1)
plot(x, pnorm(x, mean=0, sd=1),  type="l")
```



Student's t Distribution

- $X \sim t_n(\mu, \sigma^2)$

Student's t Distribution

- $X \sim t_n(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $n \in \mathbb{Z}_+$, often called degrees of freedom
- $p(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sigma\sqrt{n\pi}} \left(1 + \frac{1}{n}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{n+1}{2}}$
- Common: $t_n \sim \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$

Student's t Distribution

- $X \sim t_n(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $n \in \mathbb{Z}_+$, often called degrees of freedom
- $p(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sigma\sqrt{n\pi}} \left(1 + \frac{1}{n}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{n+1}{2}}$
- Common: $t_n \sim \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$
- $t_1 \sim$ *Cauchy* (a very dangerous distribution)

Student's t Distribution

- $X \sim t_n(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $n \in \mathbb{Z}_+$, often called degrees of freedom
- $p(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sigma\sqrt{n\pi}} \left(1 + \frac{1}{n}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{n+1}{2}}$
- Common: $t_n \sim \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$
- $t_1 \sim$ *Cauchy* (a very dangerous distribution)
- $t_{large} \sim N(0, 1)$

Student's t Distribution

- $X \sim t_n(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $n \in \mathbb{Z}_+$, often called degrees of freedom
- $p(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sigma\sqrt{n\pi}} \left(1 + \frac{1}{n}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{n+1}{2}}$
- Common: $t_n \sim \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$
- $t_1 \sim$ *Cauchy* (a very dangerous distribution)
- $t_{large} \sim N(0,1)$
- If $X \sim N(0,1)$, $Y \sim \chi^2_n$, $X \perp\!\!\!\perp Y$, then $\frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$

Student's t Distribution

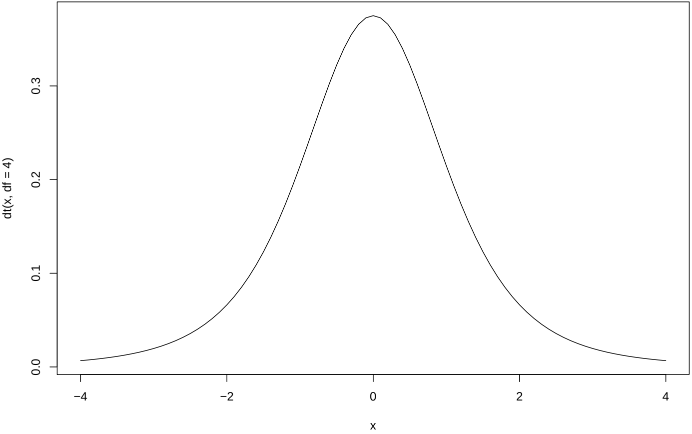
- $X \sim t_n(\mu, \sigma^2)$
- $x \in \mathbb{R}$
- $n \in \mathbb{Z}_+$, often called degrees of freedom
- $p(x) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sigma\sqrt{n\pi}} \left(1 + \frac{1}{n}\left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{n+1}{2}}$
- Common: $t_n \sim \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$
- $t_1 \sim \mathit{Cauchy}$ (a very dangerous distribution)
- $t_{large} \sim N(0,1)$
- If $X \sim N(0,1)$, $Y \sim \chi^2_n$, $X \perp\!\!\!\perp Y$, then $\frac{X}{\sqrt{\frac{Y}{n}}} \sim t_n$

Political examples

- Finite sample/unknown variance distributions
- robust estimation

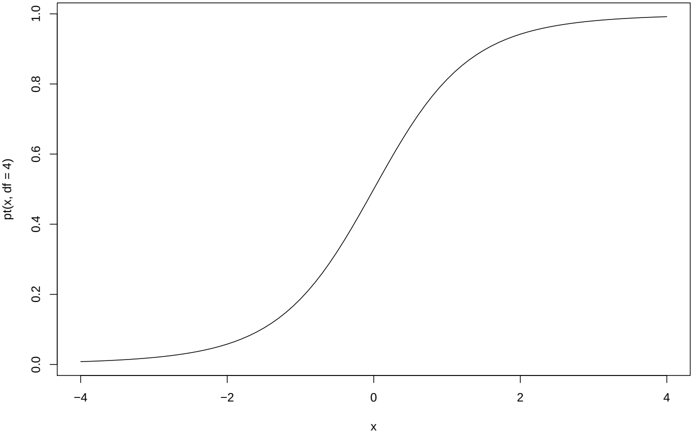
PDF for Student-t with df=4

```
x<-seq(-4, 4, by=.1)
plot(x, dt(x, df=4),  type="l")
```



CDF for Student-t with df=4

```
x<-seq(-4, 4, by=.1)
plot(x, pt(x, df=4),  type="l")
```



Other distributions you may encounter

- Logistic distribution
- F(isher's) distribution
- Gamma distribution
- Laplace distribution
- Weibull disribution
- Log-normal distribution
- Pareto distribution
- Dirichlet distribution

Joint distributions

- Often, we are interested in two or more random variables defined on the same sample space.
  The distribution of these variables is called a **joint distribution**.
- Joint distributions can be made up of any combination of discrete and continuous random variables.

Example

- Suppose we are interested in the outcomes of flipping a coin and rolling a 6-sided die at the same time.
- The sample space for this process contains 12 elements:

$$\{h1, h2, h3, h4, h5, h6, t1, t2, t3, t4, t5, t6\}$$

- We can define two random variables X and Y such that $X = 1$ if heads and $X = 0$ if tails, while Y equals the number on the die.
- We can then make statements about the joint distribution of X and Y.

Joint discret random variables

- If both X and Y are discrete, their joint probability mass function assigns probabilities to each pair of outcomes

$$p(x, y) = \Pr(X = x, Y = y)$$

- Again, $p(x, y) \in [0, 1]$ and $\sum \sum p(x, y) = 1$.

Marginal pmf

- If we are interested in the marginal probability of one of the two variables (ignoring information about the other variable), we can obtain the marginal pmf by summing across the variable that we don't care about:

$$p_X(x) = \sum_i p(x, y_i)$$

Conditional pmf

- We can also calculate the conditional pmf for one variable, holding the other variable fixed.
- Recalling from the previous lecture that $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$, we can write the conditional pmf as

$$p_{Y|X}(y|x) = \frac{p(x,y)}{p_X(x)}, \quad p_X(x) > 0$$

Joint continuous random variables

- If both X and Y are continuous, their joint probability density function defines their distribution:

$$\Pr((X, Y) \in A) = \iint_A f(x, y) dxdy$$

- Likewise, $f(x, y) \geq 0$ and $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dxdy = 1$.

Marginal pdf

- Instead of summing, we obtain the marginal probability density function by integrating out one of the variables:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Conditional pdf

- Finally, we can write the conditional pdf as

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}, \quad f_X(x) > 0$$

Expectations and moments

- We often want to summarize some characteristics of the distribution of a random variable.
- The most important summary is the expectation (or expected value, or mean), in which the possible values of a random variable are weighted by their probabilities.

Expecation of Discrete Random Variable

- The expected value of a discrete random variable Y is

$$E(Y) = \sum_y y p(y)$$

- In words, it is the weighted average of the possible values y can take on, weighted by the probability that y occurs.
- It is not necessarily the number we would expect Y to take on, but rather the average value of Y after a large number of repetitions of an experiment.

Example

- For a fair die,

$$E(Y) = \sum_{y=1}^{6} yp(y) = \frac{1}{6}\sum_{y=1}^{6} y = 7/2$$

- We would never expect the result of a rolled die to be $7/2$, but that would be the average over a large number of rolls of the die.

Expectation of a Continuous Random Variable

- The expected value of a continuous random variable is similar in concept to that of the discrete random variable, except that instead of summing using probabilities as weights, we integrate using the density to weight.
- Hence, the expected value of the continuous variable Y is defined by

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy$$

Example

Find $E(Y)$ for $f(y) = \frac{1}{1.5}, \quad 0 < y < 1.5.$

$$E(Y) = \int_0^{1.5} \frac{1}{1.5} y dy = \frac{1}{3} y^2 \Big|_0^{1.5} = .75$$

Expected Value of any probability function

1. Discrete: $E[g(Y)] = \sum_y g(y)p(y)$
2. Continuous: $E[g(Y)] = \int_{-\infty}^{\infty} g(y)f(y)dy$